

Should we have a right to refuse diagnostics and treatment planning by artificial intelligence?

Abstract

Should we be allowed to refuse any involvement of artificial intelligence (AI) technology in diagnosis and treatment planning? This is the relevant question posed by Ploug and Holm in a recent article in *Medicine, Health Care and Philosophy*. In this article, I adhere to their conclusions, but not necessarily to the rationale that supports them. First, I argue that the idea that we should recognize this right on the basis of a rational interest defence is not plausible, unless we are willing to judge each patient's ideology or religion. Instead, I consider that the right must be recognized by virtue of values such as social pluralism or individual autonomy. Second, I point out that the scope of such a right should be limited at least under three circumstances: 1) if it is against a physician's obligation to not cause unnecessary harm to a patient or to not provide futile treatment, 2) in cases where the costs of implementing this right are too high, or 3) if recognizing the right would deprive other patients of their own rights to adequate health care.

Keywords: Artificial Intelligence, right to refuse treatment, health care, patients autonomy

1. Introduction

In July 2019, *Medicine, Health Care and Philosophy* published an extraordinarily interesting article. Thomas Ploug and Søren Holm (2019) argued the need to protect the right to refuse diagnostics and treatment planning by artificial intelligence (AI). Nevertheless, the authors showed the possibility of distinguishing between a strong version of this right, which would allow the holder to refuse any involvement of AI technology in diagnosis and treatment planning, and a weak version, which would only allow recognition of the claim to physician involvement in the diagnostic and treatment planning process. The authors seemed to favour the strong version of the right, albeit with

limitations, when patients' objections are 'based on rational concerns about the systemic effects of AI use'.

In this article, I adhere to their conclusions, but not necessarily to the rationale that supports them. Instead, I criticize some of the weaknesses I found in the authors' arguments and provide some alternative arguments that might serve better to support their proposals. To this purpose, I will start by stating that it is not necessary to introduce a discussion on the weak version of the right, but on its extension. As Ploug and Holm correctly state, the right as such has been clearly recognized by the current European Union (EU) legal framework, even though we are yet to define its boundaries.

Instead, I will focus on the idea that we must adopt the strong version of the right to refuse diagnostics and treatment planning by AI, but subject to severe restrictions. To that end, I will separate myself substantially from Ploug and Holm's argumentation. First, I will argue that the idea that we cannot root this right based on a rational interest defence. I will show that this is not plausible, unless we are willing to judge each patient's ideology or religion and this is against fundamental principles included both in the Charter of Fundamental Rights of the European Union and in the General Data Protection Regulation. Instead, I will argue that the proposed right must be connected with values such as social diversity or individual autonomy and responsibility. Afterward, I will point out that the scope of such a right should be limited at least under three circumstances: 1) if it is against a physician's obligation to not cause unnecessary harm to a patient or to not provide futile treatment, 2) in cases where the costs of implementing this right are too high, or 3) if recognizing the right would deprive other patients of their rights to adequate health care.

2. The current EU legal framework: A weak version of the right to refuse diagnostics and treatment planning by AI in the GDPR

The EU legal framework on the application of AI to human health is resolved by Article 22(1) of Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation, thereafter GDPR), which reads *'The data subject shall have the right not to be subject to a decision based solely on automated processing,*

including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her.'

Obviously, the wording of this clause clearly indicates that it is not possible to use AI if there is no human element involved in the decision-making process. Therefore, we can definitely hold that the weak version of the right invoked by Ploug and Holm has already been recognised by the EU law. However, they are perfectly right when they point out that its scope is yet to be defined. Indeed, the article does not make explicit is the degree of intervention that must be considered necessary to conclude that the requirement is covered (Mitchell and Ploem 2018). Therefore, it is particularly important to highlight the statement made by the Article 29 Data Protection Working Party, an advisory body comprising a representative from the data protection authority of each EU member state, which played a prominent role in terms of interpretation of the Regulation until it was replaced by the European Data Protection Board (EDPB) under the GDPR. In 2017, the Party clarified the scope of the prohibition by stating that:

'The controller cannot avoid the Article 22 provisions by fabricating human involvement. For example, if someone routinely applies automatically generated profiles to individuals without any actual influence on the result, this would still be a decision based solely on automated processing. To qualify as human intervention, the controller must ensure that any oversight of the decision is meaningful, rather than just a token gesture. It should be carried out by someone who has the authority and competence to change the decision. As part of the analysis, they should consider all the available input and output data.'

Therefore, the legal framework is clear in one part: there is no room for solely automated decision-making, at least in the EU zone (Dreyer and Shulz 2019). However, the degree of concrete involvement of physicians in the final decision and the tools implemented to guarantee that this not become a mere tramite remains unclear. It is hard to see at the present moment how could we avoid that physicians routinely adopt the recommendations made by artificial intelligence due to defensive medicine considerations, for instance. The implementation of the weak version of the right and its concretisation includes a wide range of options and we have not faced this issue yet. Thus, there is an urgent need for discussion on this essential point and Ploug and Hold are perfectly right when they claim for it, since social concerns must play a fundamental role

in the decisions made. Hopefully, this will bring us a better agreement on the concrete degree of human involvement that the weak version of the right involves and the best ways to guarantee it.

Anyway, we could at least conclude that at the present moment, the weak version of the right to refuse diagnostics and treatment planning by AI, that is, the ‘claim to physician involvement in the diagnostic and treatment planning process’, has been endorsed by the EU regulation (Wachter et al. 2017). However, this does not at all mean that a strong version of the right, that is, the right ‘to refuse any involvement of AI technology in diagnosis and treatment planning’, is against the EU regulation. Indeed, in the next section I will argue that such a strong version of the right works well with some of the values that are widely accepted in the EU context, and thus there are some good reasons to support it.

3. The argument for the recognition of the strong version of the right

One of the parts I found most disturbing in the article by Ploug and Holm is that in the section entitled ‘Rational concerns and dystopies’ they defend the idea that the strong version of the right must be based on the patient’s rational fears and concerns. In fact, the authors make a great effort to demonstrate that if a patient raises an objection to the use of AI for those purposes on the grounds of a possible undesirable societal effect, then we should respect the patient’s claim and recognize their right to refuse diagnostics and treatment planning by AI. In my opinion, this is an unfortunate argumentation, as it concedes, in the negative sense that, if there is no rational explanation of the reasons for refusing AI, then the strong version of the right does not apply. This implies assuming the need to situate our focus on the reasonableness of a request, which means questioning the rationality of an ideology or a faith, an attitude that violates Article 21 of the Charter of Fundamental Rights of the European Union (CFR). Indeed, this is not a typical course of action, of course. Take Jehovah’s Witnesses case, for example. Do we really protect their right to refuse certain treatments on the basis of the rationality of their beliefs? In the answer to this question is the reason for my rejection of the theses of the authors of the article I now criticize (Petrini 2014).

Of course, I am not holding here that the right to refuse diagnostics and treatment planning by artificial intelligence must be considered as a part of the general right to

refuse treatment. I think that Ploug and Hold argue in a very convincing manner that both rights are different. Instead, I hold that the principles that refrain us from judging the ideology or religious beliefs that support refusing a treatment should also apply to the right that the authors of the paper that I am commenting are describing. Furthermore, I consider that there are no good reasons to oblige patients to declare the reasons why they are opposing to the use of AI in the decision process, provided that the conditions I mentioned in the introduction and will explore in the following sections apply. If this is the case, then it would only be the patient who would suffer the consequences of his or her negative. Therefore, I cannot see any strong reason to oblige him or her to reveal any kind of information about his or her ideology. Moreover, that would be contrary to the principle of data minimization, an essential ethical principle that has been incorporated to the European Union's General Data Protection Regulation (GDPR). This principle means that data processing should only use as much data as is required to successfully accomplish a given task. Provided that we can base the right to refuse diagnostics and treatment planning by artificial intelligence on reasons other than the rationality of a belief (as I will hold immediately), I do not think that we have any reason to oblige patients to reveal these very sensitive personal data. Instead, if the conditions mentioned apply, then the reasons that guide the patients' decision would be totally irrelevant, since the right would not be applicable.

Rather, I believe that we must opt for the strong version of the right based on value pluralism and the patient's autonomy and responsibility. Value pluralism means that 'people's views diverge about a range of fundamental questions, political, ethical and religious. This diversity appears to be inevitable and irresolvable. It is not possible to determine a single correct view or set of values (Turner 2004). As a consequence, negotiation, tolerance and compromise are necessary' (Wilkinson and Savulescu, 2018). Indeed, this value has been embedded in the EU Charter of Fundamental Rights in its Preamble ('The Union contributes to the preservation and to the development of these common values while respecting the diversity of the cultures and traditions of the peoples of Europe') and plays a key role in EU societies at present.

My second argument relies on the concept of autonomy, a concept that is certainly mentioned in the paper by Ploug and Hold, but in a quite different sense. They place autonomy in the basis of our acting as rational beings. In my view, autonomy refers here to the capacity of the patients to make their own decisions according to their principles

and values. Indeed, I think that respect of patient autonomy is guaranteed under Article 8 of the European Convention on Human Rights, which protects the right to private and family life. Thus, it serves as an excellent root to the right to refuse AI intervention in health care. This is probably due to the fact that I think that autonomy must be understood not only as a right to refuse a treatment, but to make decisions on the whole treatment process, as autonomy is rooted in the importance of self-government and freedom to live according to one's goals (Varelius 2006; Hartzband and Groopman 2009).

Therefore, I believe that the strong version of the right we are considering is directly connected with basic values such as patient autonomy and value pluralism and therefore it must be fully accepted in the EU context. Indeed, the focus should be on the reasons we could oppose or at least request the restrictive use of a right that is directly linked to these fundamental principles and values. What could be the reasons for defining boundaries to the right to refuse diagnostics and treatment planning by AI in its strong version? In my opinion, there are two: the need to reconcile this right with respect to physicians' ethical concerns and the costs it might involve for health care systems. I will analyse both in the next sections.

4. The argument of physicians' right to make an informed decision

First, one might oppose the right we are considering by stating that physicians are meant to have a say in the diagnostic and treatment procedures used in the development of their work. However, if we do not adopt a paternalistic approach to medicine, I doubt that this statement involves a general right for physicians to make decisions without considering the patient's values and interests. For example, in the case of Jehovah's Witnesses, we concede to such patients the right to decide on how surgery should be performed, not only the right to decide whether or not they want to undergo it. Thus, it does not seem reasonable to consider that patients cannot decide on the diagnostic and treatment tools and the possibility of avoiding AI for these purposes. However, I think that this general right only applies if this does not yield as a consequence a violation of the physician's right to not act against globally recognized medical ethical principles, such as non-maleficence or beneficence, for example (Macklin 2003). And this might certainly happen under some circumstances if we recognize a strong version of the right.

Indeed, there are cases when a physician might be unsure on whether a concrete treatment may be futile or even harmful to a patient. Imagine, for example, that a patient with cancer requests chemotherapy, but the oncologist does not know if this could be really effective in this concrete case. Under these circumstances, AI might be the only means of making a decision about it. If the patient exercises the strong version of the right, physicians would have to face a situation in which they might infringe their ethical duties: they might finally act without knowing if the intervention will not cause harm or death, not to mention the futile use of public resources, even though it would be possible to solve this dilemma using AI tools. In my opinion, cases such this show that the right to veto could undermine physicians' right to use the most accurate resources available to ensure that they are not disregarding the essential ethical principles in health care I have mentioned earlier. Indeed, physician refusal to provide futile or harmful care is supported by the ethical principle of non-maleficence, which seems particularly relevant in the situation described (Luce 1995).

Ploug and Holm might point out that this same happens in the case of Jehovah's Witnesses and yet physicians have a duty to proceed with the alternative treatment, but I think that both scenarios are not at all the same. In the case of Jehovah's Witnesses, physicians are forced to choose an alternative treatment that, in any case, would always work better than no treatment at all or death. In the case described, it might perfectly happen that the treatment would help the patient face their cancer, but it might also cause unnecessary pain to the patient. Thus, physicians forced to provide treatment without using AI would be aware that they could be causing an avoidable harm, but they would also know that if they were to not provide the treatment, they could flout the beneficence principle. The question of certitude is key in such situations, but this is precisely what the patient opting for the right would be stealing from physicians, and I think this is unfair to physicians. Moreover, I think this is a misunderstanding of the informed consent framework (Paris 2010).

Therefore, we must conclude that the right of patients to not use AI in decisions about their treatment cannot be extended to the point of forcing doctors to act against commonly accepted medical ethical principles. This could be expressed either by establishing this circumstance as a limit on the exercising of the right, or by accepting the right to conscientious objection from the health professional who is to provide the treatment. In my opinion, it seems more reasonable to adopt the first option, because if

the principle of non-maleficence is a basic principle in medical ethics, we should not think that its respect implies the need to invoke the right to conscientious objection. Furthermore, as the principles at stake are universally accepted, would it make any sense to finally put into practice a treatment that could be futile or harmful only because the patient manages to find a doctor who does not mind carrying it out? In my opinion, such a physician would be flouting the principle of non-maleficence on the basis of the alibi provided by the principle of patient autonomy.

To sum up, I consider that respect of health care workers' principles and values is a strong enough reason to conclude that the right to refuse diagnostics and treatment planning by AI cannot be an unlimited right. Indeed, it seems reasonable to think that if AI can determine whether a treatment will serve, or instead cause harm or be futile, it is the obligation of a physician to make good use of it, due to the prevalence of the non-maleficence principle, which overrides the autonomy of the patient. I concede that Ploug and Holm might be right on the idea that in the future some physicians "may come to be biased toward the decisions made by the IA technology and less sensitive to the particular preference and interests of the individual". However, if this were the case they would not be practising medicine according to the goals and standards of their profession and, thus, this type of attitudes could never become a legitimate boundary to the right we are discussing now.

5. Health care system sustainability and the rights of other patients

The second factor when considering the limits of the right to refuse diagnostics and treatment planning by AI are those that derive from the costs that the recognition of this right could cause to the health care system. Ploug and Holm address this issue in a sensible way in their paper and I cannot but adhere to what they state, even though I would like to make some remarks to it. I do not share their belief that in some circumstances, "allowing some patients to refuse AI involvement (...) might lead to cost savings because patients who are strongly opposed to AI would avoid seeking health care until their conditions have progressed to a serious state", I dare say that, in those cases, the use of AI in a previous state could lead to a better diagnosis and a more effective treatment, rendering the costly treatment that we usually have to administrate to a serious state unnecessary. To the contrary, I suspect that recognizing the strong version of the

right will probably reduce the savings that the implementation of AI in health care might bring. However, I do not think that this fact, even if confirmed by evidence, should play a definitive role in order to oppose to the recognition of the right. It is very common, in fact, for the exercise of a right derived from patient autonomy to harm public health. This happens, for example, if we accept that a patient rejects an optimal treatment, opting instead for another that will ultimately lead to higher public health costs.

However, we usually accept this result based on the defence of principles such as the need to respect the plurality of values in non-uniform societies or the importance of respecting each person's life plans. Thus, for example, no patient is forced to undergo a kidney transplant even though the alternative (that is, long-term dialysis) is much more expensive for the system. Nor has compulsory vaccination been introduced against influenza, even though this substantially increases healthcare costs. Nor, of course, are patients penalized in general for not strictly following the recommended treatment, even though this may lead to relapses and higher costs. Moreover, there are strong reasons that support such policies (Howard 2008; Schmidt 2007). I therefore understand that the concept of increased public health costs should not serve to veto, in general, the strongest version of the law we are analysing.

However, I believe that there are exceptions to this general rule. If a treatment is particularly costly, for example, it should not be administered without first having recourse to the advice of the AI if the efficiency of the corresponding predictive algorithm had been demonstrated. This usually happens in health systems, which set specific indexes for decision-making on financing treatments (which as in the UK happens with the incremental cost-effectiveness ratio [ICER]) (Nikolentzos et al 2008). I believe that this type of threshold could perfectly well be applied even in the case of recognising this right, setting objective limits to its exercise.

However, I believe that the main objection against the strong version of the right we are analysing comes from other types of situations. More specifically, it is necessary to consider cases in which the exercising of this right would cause obvious harm to third parties, who would be deprived of adequate care as a consequence. This might be better understood through an example. Imagine that eight people want to gain access to very expensive treatment. Furthermore, the statistics show that only half of the patients with that concrete condition respond to the treatment in a minimally reasonable way, following the prevailing criteria for allocating resources in that health system. Interestingly, there is

an algorithm capable of precisely guessing which of these eight people will benefit from the treatment at a reasonable cost and which will not. However, four of them refuse to have the AI used in the analysis of their specific case.

Imagine now that the AI is used for the other four and that the algorithm determines that two of them are not treatable under the underlying cost conditions. This means that there are six people left who are likely to enter into the final selection of the four candidates. If we believe that the appeal to the right should not lead to any discrimination against those who exercise it, it would be logical to draw lots among the remaining six. Thus, fortune would decide impartially who will and will not be treated. However, statistically, this would imply that at least one person capable of healing would be excluded and one for whom treatment is futile would be treated.

In my view, however, this final distribution of resources would be absurd. The logical approach would be to administer treatment to the two people for whom the AI has made an encouraging prognosis and to circumvent the other two candidates among the four who want to exercise their right to not have these mechanisms used to decide on their treatment. The opposite would be to arrive at an inefficient and unfair result based on personal ideology. However, if this is the case, then it is clear that the right we are talking about has to be limited based on the costs for the health system, the need to optimise resource allocation, but above all, on the right of third parties to access efficient treatment. It could, of course, be pointed out that the case I have put forward is exceptional and should not serve as a rule. I do not think that is true. It is a case that arises every time there is a drug shortage, and we must design a system for allocating scarce resources among patients who are likely to take advantage of it (or not). In my opinion, if AI were able to suggest an efficient form of allocation, we should not allow the right to refuse treatment planning by AI to deny scarce health resources to patients who are able to benefit from it.

References

1. Ploug T, Holm S. The right to refuse diagnostics and treatment planning by artificial intelligence. *Medicine, Health Care and Philosophy*, 2019. Jul 29. doi: 10.1007/s11019-019-09912-8. [Epub ahead of print]
2. Mitchell C, Ploem C. Legal challenges for the implementation of advanced clinical digital decision support systems in Europe. *J Clin Transl Res*. 2018;3(Suppl 3):424–430. Published 2018 Aug 18.
3. Dreyer S, Schulz W, The General Data Protection Regulation and Automated Decision-making: Will it deliver? Potentials and limitations in ensuring the rights and freedoms of individuals, groups and society as a whole. Working Paper. January 2019 Bertelsmann Stiftung. Available from: <https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/GDPR.pdf>. Accessed 11 August 2019
4. Turner L, Bioethics in pluralistic societies. *Medicine, Health Care and Philosophy*, 2004; 7(2), 201–208.
- 5 Art. 22 GDPR. 2018. Automated Individual Decision-Making, Including Profiling|General Data Protection Regulation (GDPR). General Data Protection Regulation (GDPR). <https://gdpr-info.eu/art-22-gdpr/>. Accessed 11 August 2019
- 6 Article 29 Data Protection Working Party. 2018. ARTICLE29 Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679 (wp251rev.01). https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053. Accessed 20 July 2019
7. Wachter S, Mittelstadt B, and Floridi L, Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation. *International Data Privacy Law*, 2017, 7(2):76–99.
8. Charter of Fundamental Rights of the European Union (CFR), 2012/C 326/02. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:12012P/TXT>. Accessed 20 July 2019
9. Petrini C. Ethical and legal aspects of refusal of blood transfusions by Jehovah's Witnesses, with particular reference to Italy. *Blood Transfus*. 2014;12 Suppl 1(Suppl 1):s395–s401. doi:10.2450/2013.0017-13

10. Wilkinson D, Savulescu J, Cost-equivalence and Pluralism in Publicly-funded Health-care Systems. *Health Care Analysis*, 2018, Volume 26, Number 4, Page 287
11. Howard BM, First, Do Not Punish: Individual Incentives in Health Policy, *Virtual Mentor*. 2008;10(11):719-723. doi: 10.1001/virtualmentor.2008.10.11.conl1-0811.
12. Schmidt H. Patients' charters and health responsibilities. *BMJ*. 2007;335(7631):1188
13. European Convention on Human Rights. https://www.echr.coe.int/Documents/Convention_ENG.pdf. Accessed 20 August 2019
14. Varelius J. The value of autonomy in medical ethics. *Med Health Care Philos*. 2006;9(3):377-788.
15. Macklin R. Applying the four principles. *Journal of Medical Ethics* 2003;29:275-280
16. Hartzband P, Groopman J. Keeping the patient in the equation—Humanism and health care reform. *New England Journal of Medicine* 2009, 361: 554–555
17. Luce JM, Physicians do not have a responsibility to provide futile or unreasonable care if a patient or family insists. *Crit Care Med*. 1995;23:760–766
18. Paris JJ, Autonomy does not confer sovereignty on the patient: A commentary on the Golubchuk case. *American Journal of Bioethics*, 2010: 10(3), 54–56
19. Nikolentzos A, Nolte E, Mays N, Paying for (expensive) drugs in the statutory system: an overview of experiences in 13 countries. London: London School of Hygiene & Tropical Medicine. 2008. Available from: http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_089990. Accessed 20 August 2019.

Acknowledgements:



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039